

Author copy. Publication citation:

Park, H. J., Turetsky K.M., Dahl, J., Pasek, M. H., Germano, A. L., Harper, J. O., Purdie-Greenaway, V., Cohen, G. L., & Cook, J. E. (2023). Investigating Cortisol in a STEM Classroom: The Association Between Cortisol and Academic Performance. *Personality and Social Psychology Bulletin*. doi.org/10.1177/01461672231188277

Investigating Cortisol in a STEM Classroom:

The Association Between Cortisol and Academic Performance

Hyun Joon Park

Connecticut College

Kate M. Turetsky

Barnard College, Columbia University

Julia L. Dahl

U.S. Department of Defense

Michael H. Pasek

University of Illinois Chicago

Adriana L. Germano

Columbia University

Jackson O. Harper

Indiana University at Bloomington

Valerie Purdie-Greenaway

Columbia University

Geoffrey L. Cohen

Stanford University

Jonathan E. Cook

The Pennsylvania State University

Current word count: 9982

Author Note

Hyun Joon Park, Department of Psychology, Connecticut College; Kate M. Turetsky, Department of Psychology, Barnard College, Columbia University; Julia L. Dahl, U.S. Department of Defense; Michael H. Pasek, Department of Psychology, University of Illinois Chicago; Adriana L. Germano, Business School, Columbia University; Jackson O. Harper, Department of Information and Library Science, Indiana University at Bloomington; Valerie Purdie-Greenaway, Department of Psychology, Columbia University; Geoffrey L. Cohen, Department of Psychology and Graduate School of Education, Stanford University; Jonathan E. Cook, Department of Psychology, The Pennsylvania State University.

Support for this research was provided by grants from National Science Foundation (DRL-1109548, DGE-1144155, DGE-1255832).

Correspondence concerning this article should be addressed to Hyun Joon Park, Department of Psychology, Connecticut College, New London, CT 06320, USA. Email: hpark3@conncoll.edu or Jonathan E. Cook, Department of Psychology, The Pennsylvania State University, University Park, PA 16802, USA. Email: jonathan.emdin.cook@gmail.com

This paper is the accepted version at the Personality and Social Psychology Bulletin. This paper is not the copy of record and may not exactly replicate the final published authoritative document. The final article is available at: [doi.org/10.1177/01461672231188277]

Abstract

STEM education can be stressful, but uncertainty exists about (1) whether stressful academic settings elevate cortisol, particularly among students from underrepresented racial/ethnic groups, and (2) whether cortisol responses are associated with academic performance. In four classes around the first exam in a gateway college STEM course, we investigated participants' ($N = 271$) cortisol levels as a function of race/ethnicity and tested whether cortisol responses predicted students' performance. Regardless of race/ethnicity, students' cortisol, on average, *declined* from the beginning to the end of each class and across the four classes. Among underrepresented minority (URM) students, higher cortisol responses predicted better performance and a lower likelihood of dropping the course. Among non-URM students, there were no such associations. For URM students, lower cortisol responses may have indicated disengagement, whereas higher cortisol responses may have indicated striving. The implication of cortisol responses can depend on how members of a group experience an environment.

Keywords: cortisol, URM students, academic performance, STEM, stress

Investigating Cortisol in a STEM Classroom:

The Association Between Cortisol and Academic Performance

Science, technology, engineering, and mathematics (STEM) education is important for training society's next generation of skilled workers and innovators and provides opportunities for upward social mobility (Carnevale et al., 2015). Thus, it is critical to understand factors that can hinder or help students' performance and persistence in STEM education. Elevated stress is one factor that may be a potent hindrance, given that STEM courses can be rigorous and technically challenging. However, students may not always accurately report their stress (e.g., due to motivated cognition), so it would be helpful if common biomarkers like cortisol could function as a proxy. However, little is known about whether or how cortisol responds to academic stressors or whether cortisol predicts performance and persistence. Investigating these associations is critical because elevated cortisol can interfere with cognitive processes involved in learning and recall (Vogel & Schwabe, 2016), which may impair performance and undermine motivation to persist.

Stress as assessed by cortisol may be particularly useful in helping to explain why students from underrepresented minority (URM) groups (e.g., Blacks, Hispanics, and Native Americans) underperform in STEM education compared to students from non-URM groups (e.g., Whites and Asians) (National Center for Science and Engineering Statistics [NCSES], 2021). The psychological experiences of STEM URM students in higher education can be a source of stress (Rainey et al., 2018) that undermines performance and persistence (Cook et al., 2012; Turetsky et al., 2020; Hatfield et al., 2022). However, little is known about whether URM and non-URM students differ in cortisol responses and whether or how cortisol is associated with academic performance and persistence for these groups.

To address these gaps, we longitudinally investigated students' cortisol responses in a gateway college STEM class using salivary cortisol collected at the beginning and end of four classes around the first exam. We used multilevel modeling to investigate (1) whether cortisol increased during class meetings, (2) whether cortisol rose across classes as the exam approached, and (3) whether students from URM groups, who are at higher risk of leaving STEM fields (NCSES, 2021), differed from non-URM students in cortisol levels and trajectories. We then examined the association of cortisol responses around the exam period with academic performance, including whether any such association differed by students' URM status.

Below, we describe research on stress in STEM education, including the added stress from being a member of a minoritized racial/ethnic group. Then, we discuss research on cortisol and its implications for performance and retention, including potential differences as a function of URM status.

Stress in STEM Higher Education

Students often find higher education settings to be stressful, whether due to heavy workloads (APA, 2020) or frequent evaluation and negative feedback (Stetler & Guinn, 2020). Stress may be amplified for students in STEM because of the rigorousness of many STEM curricula and the importance of satisfactory grades for progressing through required STEM courses (Michel et al., 2018).

Although higher education settings can be stressful for everyone, students from URM (compared to non-URM) groups, may experience additional stress, particularly in STEM. URM students may experience *social identity threat*, a concern of being viewed stereotypically or treated negatively, because of their race/ethnicity (Steele et al., 1997). Social identity threat can be amplified in STEM fields because URM students are even more likely to be underrepresented

(Whitcomb & Singh, 2021) and people commonly associate success with intellectual giftedness (Leslie et al., 2015). Social identity threat may help explain why URM students often feel less belonging in STEM domains (Rainey et al., 2018) and are more likely to underperform or switch fields (Hatfield et al., 2022).

Although ample evidence suggests that STEM educational can be stressful, particularly for students from URM groups, less is known about how stressful academic settings affect students' cortisol levels and how cortisol responses affect students' performance. Below, we consider the function and cognitive effects of cortisol and its interpretation.

Function and Cognitive Effects of Cortisol

Cortisol is a hormone produced by the adrenal gland as a product of the hypothalamic-pituitary-adrenocortical (HPA) axis. Cortisol raises circulating glucose to mobilize energy (Sapolsky et al., 2000). Short-term elevation of cortisol helps people manage immediate situations. However, while cortisol levels are elevated, systems less involved with immediate responses (e.g., immune functioning and complex cognition) can be suppressed. Over time, long-term suppression of these systems can have detrimental effects on cognitive functioning and health (McEwen, 1998).

Cortisol levels rise in psychologically threatening situations, particularly situations that are uncontrollable and socially evaluative (Dickerson & Kemeny, 2004). For example, during the Trier Social Stress Test (TSST), participants are required to engage in a public speaking task in front of judges who give negative feedback and calculate arithmetic aloud. These tasks are judged by most people to be psychologically stressful, and the TSST reliably elevates cortisol (Dickerson & Kemeny, 2004). However, people also can report feeling stress in other situations without elevated cortisol (Loft et al., 2007).

Cortisol also increases in situations that are psychologically neutral or even positive. For example, cortisol rises when a person wakes in the morning to ready the body for the upcoming day (Kirschbaum & Hellhammer, 1989). Cortisol can also rise in exciting situations like Christmas Eve for children (Flinn, 2006), and in situations that evoke positive affect (Hoyt et al., 2016). Thus, cortisol is not tied exclusively to negative stressful events, but to any event, that necessitates additional energy to respond (e.g., Barrett, 2017).

Of relevance to the present research is the effect of cortisol on cognitive functioning, which can affect academic performance. If cortisol rises in stressful circumstances, it may be a mechanism by which stress undermines cognition. Indeed, in laboratory studies, elevated cortisol can impair cognitive functioning, including working memory (Elzinga & Roelofs, 2005), memory retrieval (de Quervain et al., 2000), and memory encoding (Payne et al., 2007). However, it is not clear if effects of cortisol on cognition generalize to the complexity of real-world performance. Moreover, moderately elevated cortisol can improve cognitive functioning, such as working memory and memory encoding (for review, see Lupien et al., 2007), evoking the Yerkes-Dodson law on the association between stress and performance (Yerkes & Dodson, 1908). This further suggests the complicated association between cortisol, cognition, and academic performance.

Taken together, although cortisol is commonly thought of as a biomarker of stress, its biological function is largely one of mobilization, and the implication of cortisol for cognitive functioning is complex. Cortisol rises in response to some stressful negative psychological experiences, but not others, and also rises in response to some situations that are neutral or positive (e.g., McEwen, 2019). Given these complexities, more research is needed to better

understand how cortisol levels respond to natural stressors over time, as well as the utility of cortisol for predicting performance.

Cortisol in Academic Settings

As cortisol affects cognition, researchers have hypothesized that cortisol may affect learning and/or academic performance in real-life settings. However, results are mixed both on whether stress from academic settings elicits cortisol responses and whether cortisol affects real-life learning and performance (see Vogel & Schwabe, 2016).

Studies examining cortisol as an *outcome* in academic settings have tested whether salivary cortisol increases in response to stressful academic events like exams. Results are mixed. While some studies find higher cortisol levels on an exam day relative to other days (Preuß et al., 2010; Schoofs et al., 2008), others have found cortisol levels to be unchanged on an exam day (Pletzer et al., 2010). Some studies even find *lower* levels of cortisol during exam periods compared to an earlier phase of the semester (e.g., Loft et al., 2007). One clue to mixed results may be whether exams are written or oral. Cortisol levels are elevated for oral exams (Preuß et al., 2010; Schoofs et al., 2008), but results are mixed for written exams (Pletzer et al., 2010; Preuß et al., 2010). Perhaps oral exams more reliably elicit cortisol responses because of their social-evaluative nature (Dickerson & Kemeny, 2004). Importantly, most studies have methodological limitations, such as small sample sizes and a small number of cortisol assessments (typically 1-2 days), which limits the ability to assess change over time.

Only a few studies have examined cortisol as a *predictor* of academic performance in real-life settings and these too have yielded mixed results. Several studies find no association between cortisol and academic performance (Pletzer et al., 2010; Schoofs et al., 2008), while three studies report that higher cortisol levels or a change in cortisol predicts lower performance.

Ng et al. (2003) found that among 11 graduate students, higher levels of cortisol assessed before a final written exam predicted lower exam scores. Jamieson et al. (2022) found that among 339 community college students in a mathematics course, an increase in cortisol from a baseline class to the day of the exam predicted worse exam performance. Heissel et al. (2018) found that among 93 elementary- and middle-school students, change in cortisol from an earlier baseline week to just before a standardized test predicted worse test performance. Interestingly, change encompassed an increase or a decrease in cortisol, the latter perhaps suggesting disengagement.

Overall, results are ambiguous as to whether stressors in academic settings reliably elicit cortisol responses and whether (or how) cortisol predicts academic performance. Variation in methodological factors likely contributes to ambiguity, including characteristics of stressors, sample sizes, and frequency of cortisol sampling. Importantly, null effects of the association between cortisol and performance in past research may reflect an unanalyzed source of variation, such as race/ethnicity, which can affect people's reactions to academic environments. More generally, past research has not always considered the possible mobilizing nature of cortisol, which we discuss below.

Cortisol's Association with Engagement

Cortisol can also be associated with a more positive interpretation, such as task engagement (Keller et al., 2011). Keller et al. (2011) randomly assigned participants to play a computer game in three conditions varying in difficulty: boredom (low difficulty), fit (moderate difficulty tailored to participant performance), or overload (high difficulty). Contrary to the expectation that those in the overload condition would have the highest cortisol response, participants in the fit condition had the highest cortisol response – and the most involvement with the game. As noted above, cortisol has also been associated with positive affect (Hoyt et al.,

2016), which is often associated with greater engagement (King et al., 2015). If positive cortisol responses can indicate engagement, perhaps they may sometimes predict better academic performance. If so, it is important to identify factors that affect cortisol's associations with different psychological states, a topic to which we turn next.

Interpreting Physiological Arousal

Research has long suggested that physiological arousal can be associated with different psychological states depending on how individuals appraise and react to their social environment (e.g., Schachter & Singer, 1962; Blascovich & Mendes, 2010). For instance, Crum and colleagues (2017) had participants complete a TSST and found increases in cortisol after the task. However, participants who had been induced earlier into a stress-is-enhancing mindset (stress improves outcomes) reported more positive affect and had higher levels of dehydroepiandrosterone-sulfate (DHEAS), an anabolic hormone with cellular protective properties (Rutkowski et al., 2014), than those induced into a stress-is-debilitating mindset (stress impedes outcomes). Similarly, among participants engaged in a stressful salary negotiation task, elevated cortisol was *positively* associated with negotiation performance for those told that physiological arousal was helpful, but *negatively* associated with performance for those told nothing about physiological arousal (Akinola et al., 2016). In another study, the association between cortisol responses and performance differed as a function of participants' anxiety levels (Mattarella-Micke, 2011). Collectively, these studies suggest that even in situations that elevate cortisol, the association between cortisol and performance may be malleable depending on how people appraise themselves within an environment. If so, it is important to identify factors that influence individuals' experiences in an academic environment. Race/ethnicity may be one such factor.

Different Experiences in Academic Settings Based on Race/Ethnicity

Students from different racial/ethnic groups often experience academic settings differently, which may affect cortisol levels and cortisol's association with performance. Below, we consider the possible implications of cortisol separately for URM and non-URM students.

Understanding Cortisol for URM Students

Cortisol as an Outcome. Because URM students may experience STEM academic environments as more unwelcoming, and this is likely to be stressful, URM students in these environments may have higher cortisol levels than non-URM students (Levy et al., 2016). We hypothesized and pre-registered this possibility (see <https://bit.ly/3EB2v5i>).

The longitudinal research design of this project also led to a consideration of time. In general, we expected that the stress of the course might cause cortisol to rise within classes and from the beginning of the semester to around the first exam. Within classes, students may have experienced increasing stress if they found the material challenging. Across classes, students may have experienced greater stress because the first exam was approaching, which would likely lead to concerns of evaluation, particularly given the importance of this course on eligibility to continue in the major and career path that students had chosen. Following the logic that URM students may experience a greater stress burden from social identity threat, we reasoned that URM students' cortisol levels would increase more during class than it did for non-URM students. We also expected that URM students' cortisol levels would rise more from the beginning of the semester to around the first exam than it would for non-URM students.

Cortisol as a Predictor of Performance. Whether or not URM students differ from others in their cortisol responses, most researchers have hypothesized a negative association between cortisol and academic performance (Levy et al., 2016). Based on this logic, our a-priori

pre-registered hypothesis was that elevated cortisol responses around the first exam of the semester would negatively affect performance and help explain the achievement gap between URM and non-URM students (see <https://bit.ly/3EB2v5i>).

However, based on a review of the literature, we considered an alternative hypothesis (not pre-registered) that cortisol responses may be positively associated with performance for URM students because they are more likely to use two strategies in the classroom to cope with social identity threat: (1) striving harder to overcome negative stereotypes and (2) disengaging to protect self-worth (James et al., 1983; Major & O'Brien, 2005). If so, among URM students, higher cortisol may represent striving and thus predict better performance, whereas lower cortisol may represent disengagement and thus predict worse performance.

Understanding Cortisol for Students from non-URM Groups

Cortisol as an Outcome. Our a-priori prediction was that non-URM students would be less likely to experience race-based social identity threat in a typical STEM classroom setting, and thus, would have lower cortisol levels and less steep increase in cortisol within- and between- classes than URM students.

Cortisol as a Predictor of Performance. Following conventional wisdom, we expected that higher cortisol responses would predict worse performance for non-URM students, as with URM students. However, because non-URM students would have less need to cope with social identity threat, we did not think it likely that there would be a positive association between cortisol responses and performance for non-URM students.

Current Study

To better understand the utility and interpretation of cortisol in an academic setting, we measured students' cortisol levels at the beginning and end of four classes from early in the

semester until shortly after the first exam in a racially diverse gateway college STEM course, Introduction to Molecular and Cellular Biology. This required course for students planning on medical or science careers has a high failure rate, suggesting its suitability for investigating cortisol responses (see S1 for details). Our research design provides insight into the longitudinal pattern of cortisol levels within classes and across classes and provides needed data on the association between cortisol responses and academic performance in an externally valid context. Moreover, the size and diversity of the class allow us to assess group differences in cortisol and its association with performance as a function of a salient social category – race/ethnicity.

Method

This research is part of an IRB approved project examining students' experiences in a large gateway college STEM class (see also Turetsky et al., 2020). The large enrollment increased the opportunity to obtain a sufficient sample of URM students and helps maximize statistical power. Because our sample size was determined by the number of students in the class, sensitivity power analyses were conducted using G*Power (Faul et al., 2007) to determine the minimal detectable effect (MDE) of a predictor in a multiple regression, controlling for covariates. Based on our final sample size of 271, $\alpha = .05$, and power of .80, results revealed a MDE of Cohen's f^2 of .028. Thus, our analyses were well-powered to detect a small effect.

Participants

A total of 552 students were enrolled in this class, of whom 328 (59%) consented to participate. Another 53 did not submit any saliva samples and were thus excluded. Those without saliva samples did not differ from those who did in demographic characteristics or class performance. Of the remaining 275, two were excluded because they dropped the class before taking any exams. Another two were excluded because their average cortisol levels were more

than three standard deviations above the mean (Dozier et al., 2006). The final sample size was 271.

Participants' mean age was 21.18 ($SD = 3.91$) (see Table 1). Women outnumbered men by about 2:1 (63.8% female, 34.7% male, 1.5% gender non-binary). Participants identified as White (34.7%), Asian (27.3%), multi-racial (18.1%), Hispanic (9.2%), Black (7.7%), Native American (0.7%), and other (2.2%). We categorized students who were *at least partly* a member of an underrepresented racial/ethnic group (i.e., Black, Hispanic, Native American) as URM ($N = 78$) and those in more traditionally represented racial/ethnic groups (i.e., White, Asian, or multi-racial White and Asian) as non-URM ($N = 193$). Results from a baseline survey at the start of the class revealed that URM students had higher levels of race-based social identity threat than non-URM students ($b = 0.36, p < .001$). See S2 – S4 for additional details about participants and race/ethnicity coding.

Procedure

We emailed all enrolled students a study description and a link to begin participation before classes started and recruited in-person at the start of the semester (see S5 – S9 for additional details about procedures).

Table 1

Participant Demographic Information

Measure	Total (%) (N = 271)	Non-URM students	URM students
		(N = 193)	(N = 78)
Age			
M (SD)	21.18 (3.91)	21.46 (4.21)	20.50 (2.94)
Gender			
Male	94 (34.7%)	69 (35.8%)	25 (32.1%)
Female	177 (65.3%)	124 (64.2%)	53 (67.9%)
Parents' education level***			
Less than high school	8 (30%)	1 (0.5%) -	7 (9.0%) +
High school diploma	28 (10.3 %)	16 (8.3 %)	12 (15.4%)
Two-year college degree	16 (5.9%)	8 (4.1%)	8 (10.3%)
Four-year college degree	45(16.6%)	30 (15.5%)	15 (19.2%)
Master's degree	64 (23.6%)	45 (23.3%)	19 (24.4%)
Ph.D. or professional degree	110 (40.6%)	93 (48.2%) +	17 (21.8%) -
Self-reported parent's household income***			
Below \$30,000	21 (7.7%)	8 (4.1%) -	13 (16.7%) +
\$30,001-\$50,000	29 (10.7%)	16 (8.3%)	13 (16.7%)
\$50,001-\$90,000	47 (17.3%)	28 (14.5%)	19 (24.3%)
\$90,001- \$150,000	60 (22.2%)	43 (22.3%)	17 (21.8%)
More than \$150,000	78 (28.8%)	66 (34.2%) +	12 (15.4%) -
Unsure	36 (13.3%)	32 (16.6%)	4 (5.1%)
Self-reported year in college			
1 st year	9 (3.3%)	8 (4.1%)	1 (1.3%)
2 nd year	148 (54.6%)	102 (52.8%)	46 (59.0%)
3 rd year	43 (15.9%)	30 (15.5%)	13 (16.7%)
4 th year	6 (2.2%)	4 (2.1%)	2 (2.6%)
5 th year	1 (0.4%)	0 (0.0%)	1 (1.3%)
Postbac ¹	64 (23.6%)	49 (25.4%)	15 (19.2%)

Note. Statistical tests compare non-URM and URM students on each demographic category. Differences by age were tested using one-way ANOVA. Differences in categorical variables were tested with chi-squared tests. +/- signs indicate that cell proportions were bigger/smaller than expected based on adjusted standardized residuals > |3|. ¹Post-bac students did not differ from other students in SES, racial composition, or performance in the class, but they were more likely to be male. The pattern of results does not change if we control for students' post-bac status.

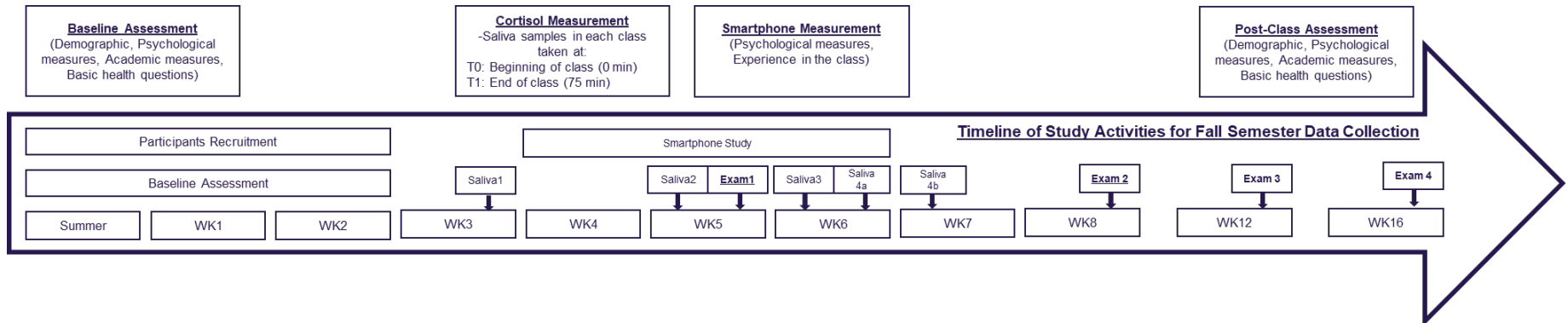
p* ≤ .05. *p* ≤ .01. ****p* ≤ .001.

Saliva samples were collected during class in Weeks 3, 5, 6, and 7 of the 16-week semester (see Figure 1). This timing was meant to incorporate students' experiences around the first midterm (Week 5). Saliva Class 1 (Week 3) provides a baseline before students perceived acute pressure from Exam 1. Saliva Class 2 (Week 5) samples were collected one class before

Exam 1, while Saliva Class 3 (Week 6) samples were collected one class after Exam 1, before students had received their grades. Saliva Class 4 (Week 6 - 7) samples were collected about a week after the exam shortly after grades were posted, and around two weeks before Exam 2.

Figure 1

Study Timeline



Note. The fourth class was rescheduled to two different days straddling Weeks 6 and 7 because of a religious holiday on the originally scheduled class day; class time was unchanged.

Participants submitted saliva samples at the beginning and end of class, which allows within- and between-class analyses. Students could attend a morning (10:10 – 11:25 AM) or afternoon (4:10 – 5:25 PM) session of each class. Given cortisol's diurnal circadian rhythm (Kirschbaum & Hellhammer, 1989), we control in all analyses for time since awakening at the moment of saliva collection to eliminate any spurious results based on time of day. Additionally controlling for class time does not change the significance and pattern of results (see S6).

Researchers handed students plastic bags as they arrived at class. The bags contained two Salivette tubes, instructions, and a short questionnaire assessing behaviors that can affect saliva (e.g., Granger et al., 2009). Participants were also emailed the night before and asked to refrain from these behaviors before class. Participants were instructed to complete the questionnaire and provide the first saliva sample as soon as they sat down and to provide the second saliva sample right before leaving class. Saliva samples were stored at or below -25°C .

The larger project included an affirmation intervention or control exercise delivered in Week 3 (Cohen et al., 2006). The intervention had a small, difficult-to-interpret effect on longitudinal cortisol patterns but did not otherwise moderate any associations of race/ethnicity or moderate the association of cortisol responses with performance. Controlling for intervention condition does not change the pattern of any results, and we do not describe the intervention further in the main text (S8 provides details).

An embedded smartphone study was conducted over 21 days during Weeks 4 – 6 among a subsample of 98 participants. The goal of the smartphone study was to capture detailed daily changes in students' psychological states around the first exam (see S9 for details). In a later section, to better understand the psychological interpretation of cortisol patterns, we explore the association between cortisol and psychological states assessed among the subsample who

completed the smartphone study.

At the end of the semester, students completed measures similar to those from baseline. Results from the post-class measures are not presented in the current study.

Measures

Primary Measures

Cortisol. Frozen saliva samples were assayed using a sensitive immunoassay (Salimetrics, State College, PA) at the Biomarker Core Lab at the Pennsylvania State University. After excluding eight samples with insufficient saliva, we had 1,745 saliva samples from 271 participants. Samples had acceptable averaged intra- and inter-assay coefficients (5.6% and 5.8%) (see S10 for details).

Race/ethnicity. Students first described their race/ethnicity with a free-response item (“How would you describe your race/ethnicity?”) and then separately selected any racial categories that applied from a pre-defined checklist (i.e., White, Black, Asian, Hispanic, Native American, multi-racial, and other). Students could provide additional information with responses of multi-racial and other. We primarily used students’ free-response descriptions of their race/ethnicity in race/ethnicity coding, but if students did not provide a free-response description or their response was unclear, we utilized checklist responses.

Final Class Score. Students took exams for this class in Weeks 5, 8, 12, and 16 (see Figure 1) with scores calculated and provided to the research team by the instructor. Our primary DV was students’ final score, which had a maximum point-value of 317 and a class average of 68% ($M = 217.00$, $SD = 54.57$). We use final score instead of final letter grades, largely to retain in the analyses 29 students who had taken some exams but dropped the class and thus, did not

receive a final letter grade. The pattern of results did not change if these students were excluded (see S11 for details).

Dropping the Course. Another primary DV was a dichotomous variable we created to indicate whether students had dropped the course based on student transcripts. Overall, 10.7% ($n = 29$) of students dropped the course, a rate that was higher among URM students (16.7%) than non-URM students (8.3%).

Exam 1 Performance. We also examined Exam 1 as the performance outcome most proximal to the saliva samples. Exam 1 had four problems, each with several short-answer questions, worth a total of 114 points with a class average of 65% ($M = 73.76$, $SD = 13.09$). Exam 2 followed a similar procedure (see S11).

Baseline Survey Covariates and Exploratory Outcomes

Social Identity Threat. The baseline survey assessed social identity threat using six items adapted from Cohen & Garcia (2005) (e.g., “I worry that people at [university] will judge me, based on what they think of my racial group”) on a scale from 1 (*not at all true*) to 5 (*very true*). Responses were averaged ($\alpha = .89$) with higher scores indicating greater identity threat.

Gender. We control for gender because cortisol responses (Hostinar et al., 2014) and performance (Eddy et al., 2017) can differ by gender. However, in our sample, students’ cortisol levels, the longitudinal pattern of cortisol, and the association between cortisol responses around the exam and performance did not differ by gender.

Participants could identify as Male, Female or Other. Four participants selected Other. Neither (1) removing these participants nor (2) creating a third gender category affected the pattern of results. Considering the small number of participants involved, and that gender was a

covariate, we grouped these participants with females for the analyses here based on their phenotypical presentation.

Socioeconomic Status (SES). SES can influence cortisol patterns (Chen et al., 2009) and performance (Reardon, 2011). As we observed a difference in SES between racial/ethnic groups, we included SES as a covariate. Students' cortisol levels, the longitudinal pattern of cortisol, and the association between cortisol responses around the exam and performance did not differ by SES.

We calculated SES based on (1) family income, assessed from 1 (\$0 – \$30,000) to 9 (\$250,000+); (2) parental education, assessed from 1 (*less than high school*) to 6 (*Ph.D. or professional degree*); and (3) perceived social status, assessed using the MacArthur Scale of Subjective Social Status (Adler et al., 2000), which has participants locate where they belong on a 10-rung ladder representing, from low to high, the income, education, and occupation of others in the U.S. These variables were standardized and averaged to form a single measure with higher scores indicating higher SES (see S12).

Transcript Covariates

Prior Academic Performance. In analyses predicting performance, we controlled for students' average GPA from 0 (F) to 4.33 (A+) in prior college biology and chemistry courses, which could affect performance in the current course, as well as cortisol levels, and the association of cortisol with performance.

Saliva Questionnaire Covariates

We controlled for participants' time since awakening at saliva collection, whether they had taken medication that can affect cortisol in the last-24 hours, and whether they had brushed their teeth in the hour before saliva collection.

Transparency and Openness

We report how we determined our sample size and all data exclusions. All data, analysis code, codebook and research materials are available at <https://bit.ly/3iaIgnp>. Materials for the larger study (i.e., baseline and smartphone study measures, intervention materials, and pre-registration) are also available: <https://bit.ly/3EB2v5i>. Data were analyzed using SPSS 26 and R (R Core Team, 2020) (see S15 for all R packages used). The overall design was pre-registered, as was the hypothesis that in the absence of intervention, at-risk students (i.e., students who may be more prone to experiencing psychological threat in academic settings, including URM students) would have higher cortisol levels on average and over time and that higher cortisol would lead to worse performance. If students' cortisol levels had differed by URM status, we had planned to test if cortisol mediated any association between URM status and performance. However, as noted below, students' cortisol levels did not differ as a function of URM status. We did not pre-register other analyses.

Results

We first present strategies for variable coding and missing data. Then, we present descriptive analyses. We then describe the longitudinal analyses investigating cortisol patterns within and between classes, including whether patterns differed by race/ethnicity. Then, we move to analyses predicting academic performance from cortisol responses around the exam and whether this association was moderated by race/ethnicity. Lastly, we present exploratory analyses to better understand observed patterns.

Data Coding Strategy

The dichotomous variable representing students' URM status was contrast coded. Students who were *at least partly* a member of a URM group (NCSES, 2021) were coded 1, and

those in more traditionally non-URM groups were coded -1. Gender was also contrast coded (male = -1, female = 1). We dummy coded the two categorical cortisol covariates (0 = No, 1 = Yes), for whether participants had (1) taken one or more medications or (2) brushed their teeth before saliva collection.

Missing Data Strategy

To avoid listwise deletion because of missing covariate data, we imputed reasonable estimates whenever possible. We took a case-by-case approach, as good approximations of missing values were available. The direction and statistical significance of results are unchanged if all cases with missing data are listwise deleted (see S16).

Descriptive Analyses and Intercorrelations

Descriptive statistics and inter-correlations of primary study variables are shown in Table 2. Students' average cortisol level across the classes was 0.24 ($SD = 0.15$), which is within the normal range among healthy adults (Aardal & Holm., 1995). Demographic factors were not associated with students' average cortisol levels; thus, our pre-registered hypothesis that URM students would have higher cortisol levels than non-URM students was not supported. With respect to students' final class scores, URM (compared to non-URM) students, women (compared to men), and lower-SES (compared to higher-SES) students had worse performance.

Table 2

Descriptive Statistics and Inter-Correlations of Primary Study Variables

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9
1. Cortisol level ¹	0.24	0.15									
2. Final scores ²	217.00	54.57	.05								
3. Gender ³	0.31	0.95	.11	-.16**							
4. Race/ethnicity ⁴	-0.42	0.91	.07	-.26***	.04						
5. SES ⁵	0.01	0.83	.002	.23***	-.01	-.39***					
6. Prior performance	3.40	0.49	-.06	.49***	-.18**	-.26***	.28***				
7. Time since awakening ⁶	4.41	2.34	-.58***	-.08	-.07	-.07	.01	-.04			
8. Medication ⁷	0.14	0.29	.14*	-.002	.09	-.03	.14*	-.14*	0.02		
9. Brush teeth ⁸	0.21	0.32	.42***	.07	.04	.07	.06	.11	-.44***	-.01	

Note. *N* = 271. ¹ Raw cortisol concentration averaged across four classes (µg/dL). ² Final class score calculated by the instructor. ³ Male = -1, Female = 1. ⁴ Non-URM = -1, URM = 1. ⁵ An average of standardized SES variables (parent’s income, parent’s education, social status). ⁶ Time since awakening averaged across classes (in hours). ⁷ Whether students took medication 24 hours before saliva collection, averaged across the four classes in which saliva samples were collected (0 = No, 1 = Yes). ⁸ Whether students had brushed their teeth an hour before the saliva collection, averaged across the four classes in which saliva samples were collected (0 = No, 1 = Yes).

p* ≤ .05. *p* ≤ .01. ****p* ≤ .001.

Primary Analyses – Longitudinal Analysis of Cortisol Within and Across Classes

To test whether students' cortisol concentrations varied within and across classes, we conducted 3-level regression analyses with using the nlme package (Pinheiro et al., 2017) available within R (R Core Team, 2020). Level 1 modeled students' within-class cortisol concentrations (i.e., time within class) and their time since awakening. Following established procedures, cortisol values were log-transformed to reduce skewness (from 3.93 to -0.03).

Level 2 modeled students' average log-transformed cortisol concentrations for each class, along with class-level cortisol covariates as predictors of the Level-1 intercept. We modeled time at Level 2 to test for cortisol patterns across classes. We began with a linear-only model and then examined possible non-linear patterns by adding quadratic and cubic terms. A model with linear, quadratic, and cubic terms fit better than either a linear-only model, $\chi^2(2, N = 271) = 16.53, p < .001$, or a linear-plus-quadratic model, $\chi^2(1, N = 271) = 14.30, p < .001$. Consequently, we retained the linear, quadratic, and cubic predictors to best capture observed cortisol patterns but used a linear-only model to test for the hypothesized increase in cortisol.

Level 3 modeled between-person demographic variables. We first include these only as predictors of the intercept to test for differences in average cortisol levels as a function of URM status, as well as gender and SES. Then, we tested whether within- and between-class cortisol patterns varied as a function of URM status and the other demographic variables by including between-person demographic variables as moderators of all time variables at Levels 1 and 2.

Intercepts at all levels were allowed to vary, as were the within-class slope and the linear slope between classes; the quadratic and cubic terms were fixed to avoid convergence errors. Models were estimated using full-information maximum likelihood and an unstructured covariance of random effects.

Below, we specify the multilevel model. The subscript t = the time cortisol was collected within each class, d = the classes when saliva samples were collected, and i = individuals.

Demographic variables in bold font were modeled as moderators in a second step after conducting the analysis with demographic variables only predicting the intercept.

$$\text{Level 1: Cortisol value}_{tdi} = \beta_{0di} + \beta_{1di}(\text{Time in class}_{tdi}) + \beta_{2di}(\text{Time since awakening}_{tdi}) + e_{tdi}$$

$$\text{Level 2: } \beta_{0di} = \gamma_{00i} + \gamma_{01i}(\text{Time across class}_{di}) + \gamma_{02i}(\text{Time across class}_{di})^2 + \gamma_{03i}(\text{Time across class}_{di})^3 + \gamma_{04i}(\text{Medication}_{di}) + \gamma_{05i}(\text{Brush teeth}_{di}) + u_{0di}$$

$$\beta_{1di} = \gamma_{10i} + u_{1di}$$

$$\beta_{2i} = \gamma_{20i}$$

$$\text{Level 3: } \gamma_{00i} = \pi_{000} + \pi_{001}(\text{Race}_i) + \pi_{002}(\text{SES}_i) + \pi_{003}(\text{Gender}_i) + r_{00i}$$

$$\gamma_{01i} = \pi_{010} + \boldsymbol{\pi}_{011}(\text{Race}_i) + \boldsymbol{\pi}_{012}(\text{SES}_i) + \boldsymbol{\pi}_{013}(\text{Gender}_i) + r_{01i}$$

$$\gamma_{02i} = \pi_{020} + \boldsymbol{\pi}_{021}(\text{Race}_i) + \boldsymbol{\pi}_{022}(\text{SES}_i) + \boldsymbol{\pi}_{023}(\text{Gender}_i)$$

$$\gamma_{03i} = \pi_{030} + \boldsymbol{\pi}_{031}(\text{Race}_i) + \boldsymbol{\pi}_{032}(\text{SES}_i) + \boldsymbol{\pi}_{033}(\text{Gender}_i)$$

$$\gamma_{04i} = \pi_{040}$$

$$\gamma_{05i} = \pi_{050}$$

$$\gamma_{10i} = \pi_{100} + \boldsymbol{\pi}_{101}(\text{Race}_i) + \boldsymbol{\pi}_{102}(\text{SES}_i) + \boldsymbol{\pi}_{103}(\text{Gender}_i) + r_{10i}$$

$$\gamma_{11i} = \pi_{110}$$

$$\gamma_{12i} = \pi_{120}$$

$$\gamma_{13i} = \pi_{130}$$

Time within classes was coded so that -0.5 = the beginning of the class and 0.5 = the end of the class. Time across class was coded so that -1.5 = Saliva Class 1 (early-semester baseline), -0.5 = Saliva Class 2 (collected one class before Exam 1), 0.5 = Saliva Class 3 (collected right after Exam 1), and 1.5 = Saliva Class 4 (i.e., shortly after students' midterm grades were posted and around two weeks before Exam 2). Time since awakening, medications, and tooth brushing were grand mean centered, as was SES.

Typical Cortisol Levels

Table 3 indicates that students' average, log-transformed within-class cortisol level was -.76. This level did not vary by race/ethnicity, gender, or SES.

Table 3

Three-Level Models of Cortisol Patterns in 271 Participants (1,745 Cortisol Samples)

Fixed effects	Estimate (<i>SE</i>)	95% CI	<i>t</i>	<i>p</i>	<i>df</i>
Level 1 (within-class)					
Average cortisol level (intercept)	-0.76*** (0.01)	[-0.78, -0.73]	-53.36	< .001	865
Time in class ¹	-0.14*** (0.01)	[-0.16, -0.13]	-15.50	< .001	865
Time since awakening ^{2,6}	-0.06*** (0.003)	[-0.07, -0.06]	-21.76	< .001	865
Level 2 (between-class)					
Class (linear) ³	0.03 (0.02)	[-0.01, 0.07]	1.64	.10	602
Class (quadratic) ⁴	0.01 (0.01)	[-0.003, 0.02]	1.48	.14	602
Class (cubic) ⁵	-0.03*** (0.01)	[-0.05, -0.02]	-3.83	< .001	602
Medication ⁶	0.02 (0.02)	[-0.02, 0.06]	0.87	.39	602
Brush teeth ⁶	0.06** (0.02)	[0.02, 0.09]	2.88	.004	602
Level 3 (between-person)					
Race/ethnicity ⁷	0.01 (0.01)	[-0.01, 0.04]	1.03	.30	267
Gender ⁸	0.01 (0.01)	[-0.01, 0.03]	0.94	.35	267
SES ^{6,9}	0.001 (0.01)	[-0.03, 0.03]	0.11	.92	267
Random effects	Variance (<i>SE</i>)	95% CI			
Level 1 (within-class)					
Intercept	0.16 (0.05)	[0.06, 0.43]			
Time in class ¹	0.18 (0.09)	[0.01, 4.17]			
Level 2 (between-class)					
Intercept	0.14 (0.01)	[0.13, 0.16]			
Time in class ¹	0.08 (0.01)	[0.07, 0.17]			
Class (linear) ³	0.04 (0.01)	[0.03, 0.06]			

¹ Time when saliva samples were collected: -0.5 = start of class, 0.5 = end of the class. ² Time since awakening coded in hours. ³ Saliva class number: -1.5 = first class, -0.5 = second class, 0.5 = third class, 1.5 = fourth class. ⁴ Class number squared, ⁵ Class number cubed. ⁶ Centered at the sample mean. ⁷ Non-URM = -1, URM = 1. ⁸ Male = -1, Female = 1. ⁹ An average of standardized SES variables (parent’s income, parent’s education level, social status).

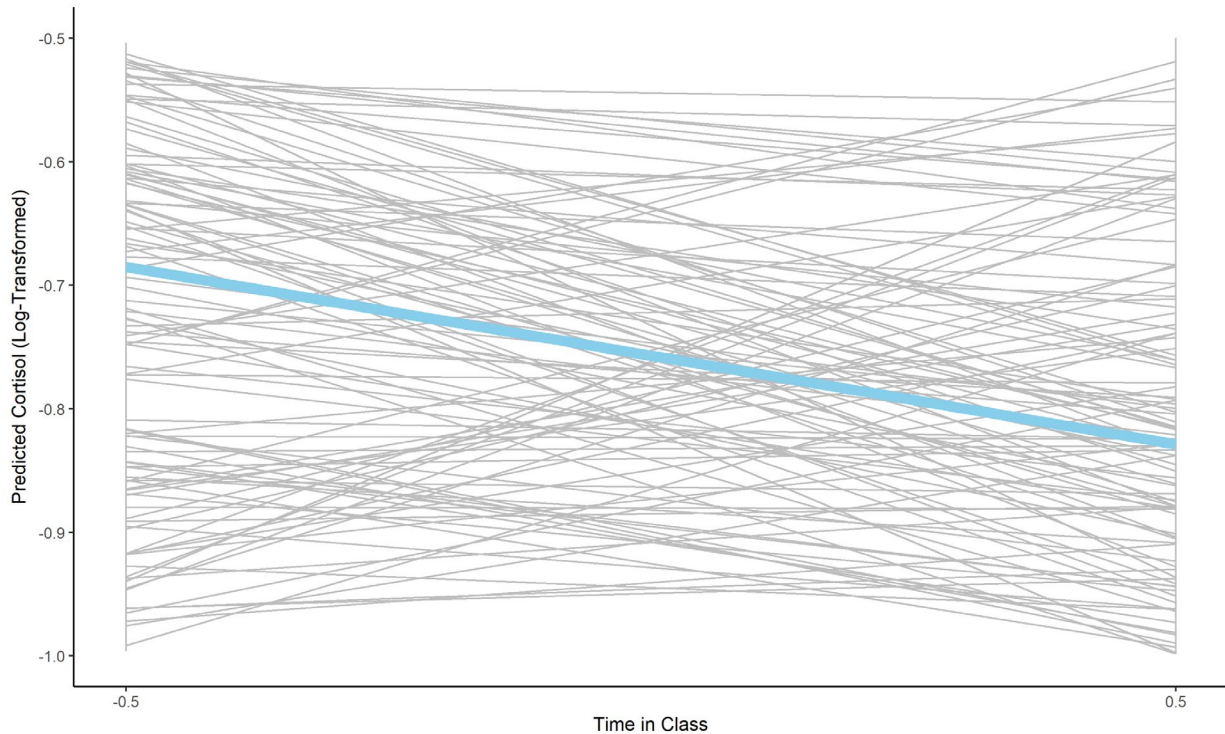
p* ≤ .05. *p* ≤ .01. ****p* ≤ .001.

Cortisol Pattern Within Class

Cortisol levels declined from the beginning to the end of class, $\beta = -0.14$, $t(865) = -15.50$, $p < .001$ (see Figure 2 and Table 3).

Figure 2

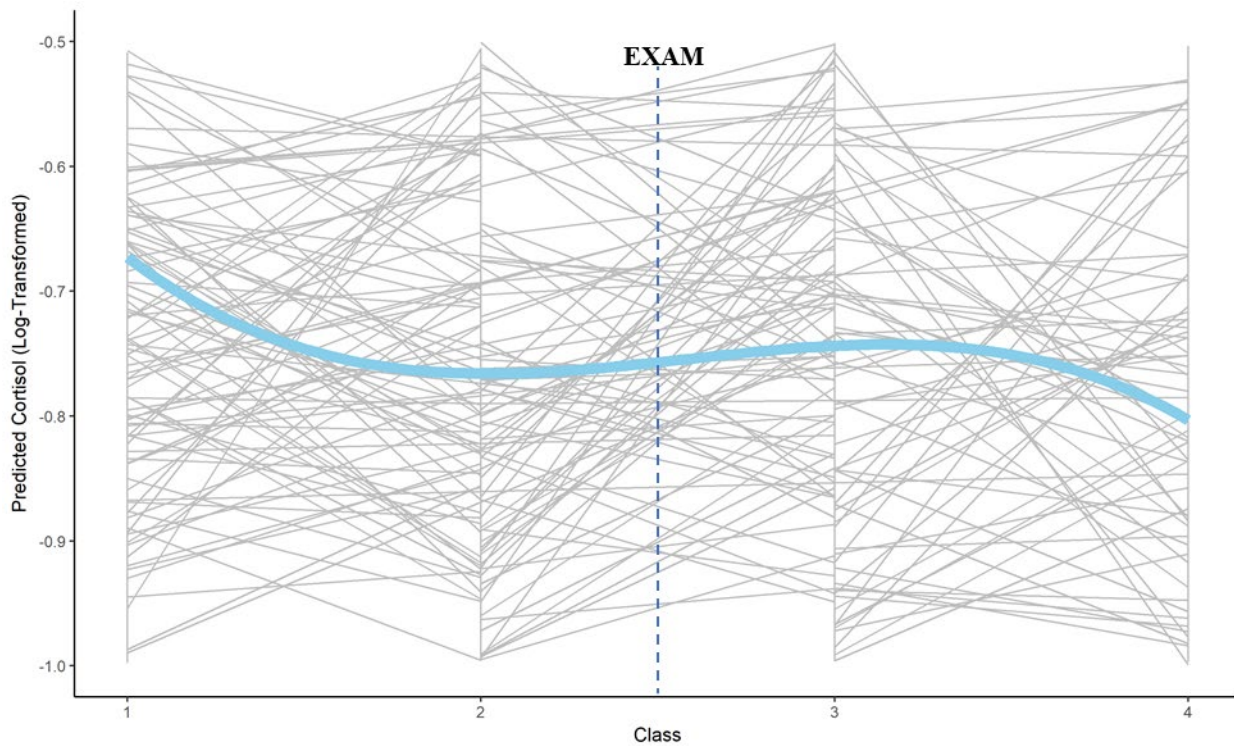
Cortisol Pattern Within Class



Note. Grey lines represent students' average within-class slope based on predicted, log-transformed cortisol values controlling for relevant covariates. The solid blue line represents students' average within-class slope across the classes (see Table 3). Saliva samples were collected at the beginning (-0.5) and end (0.5) of each 75-minute class. Class number was centered so that 0 was the theoretical average class.

Cortisol Pattern Between Classes

Cortisol levels followed a downward cubic slope across the four classes, $\gamma = -0.03$, $t(602) = -3.83$, $p < .001$ (see Figure 3 and Table 3). In a separate model with only the linear slope of time, cortisol declined across classes, $\gamma = -0.04$, $t(604) = -6.70$, $p < .001$.

Figure 3*Cortisol Pattern Across Classes*

Note. Grey lines represent students' average between-class slope of predicted cortisol values (log-transformed) averaged within classes and controlling for relevant covariates. The solid line in blue represents the average between-class slope across students (See Table 3). Class number was coded as Saliva Class 1 = -1.5, Saliva Class 2 = -0.5, Saliva Class 3 = 0.5, Saliva Class 4 = 1.5.

Interaction Between Cortisol Trajectories and Demographic Factors

To investigate whether cortisol trajectories differed by URM status, as well as gender and SES, we tested a second model that included these demographic variables as predictors of all time slopes. Adding the interaction terms as a block did not significantly improve model fit ($p = .59$), and no individual interaction coefficients were significant. We also tested whether the declining pattern of within-class cortisol trajectories changed across the semester. While there was some variability in magnitude, cortisol trajectories declined within every class (see S6).

Primary Analyses - Examining the Association Between Cortisol Responses Around the Exam and Academic Performance

Data Analysis Plan

We tested whether students' cortisol responses around the exam period, controlling for their earlier cortisol levels, predicted final performance and dropping the class. Specifically, we regressed the two performance outcomes on baseline cortisol level (Saliva Class 1) and cortisol levels around the exam (an average of Saliva Class 2 - 4), ultimately adding performance and cortisol covariates and interactions with race/ethnicity, as described below. The second cortisol predictor tests whether residual cortisol from around the exam period, controlling for cortisol levels earlier in the semester, predicts performance and persistence. For Exam 1 performance, the second cortisol predictor was cortisol levels assessed in the class before Exam 1 only (i.e., Saliva Class 2 cortisol). We also tested whether controlling for Saliva Class 1 cortisol, the association between averaged cortisol from Saliva Class 2 – 4 predicted Exam 2 performance, approximately two weeks after Saliva Class 4. This approach yielded similar results (see S17).

We used a hierarchical model-building approach. In *Step 1*, we included only the cortisol predictors, which were averaged and then log-transformed to reduce skewness. In *Step 2*, we added covariates associated with cortisol responses and/or academic performance (see Table 4) to eliminate potential confounding explanations. In *Step 3*, we tested the interaction between cortisol responses and race/ethnicity in predicting students' academic performance. We also exploratorily tested whether the interaction between cortisol responses and gender and SES significantly predicted performance. They did not (see S18).

The sample size for the primary outcomes was reduced to 234 from 271 because participants either did not submit samples at Saliva Class 1 or across Saliva Class 2-4. The

sample size is further reduced for analyses focusing on Exam 1 to the 192 students who submitted saliva samples for both Saliva Class 1 and 2 and took the first exam. The results do not change if averaged cortisol levels across four classes are used as the predictor of performance ($N = 271$) (see S19).

Cortisol and Students' Final Scores

In *Step 1*, controlling for Saliva Class 1 cortisol, students' averaged cortisol levels from later classes were not associated with their final scores, $b = 19.88$, $F(1, 231) = 1.27$, $p = .26$, 95%, $\eta_p^2 = .005$ (see Table 4 for results from all steps and 95 % CIs). Adding covariates in *Step 2*, higher cortisol responses predicted marginally *higher* final scores, $b = 32.36$, $F(1, 223) = 3.54$, $p = .061$, $\eta_p^2 = .016$. We further tested an inverted U-shaped association between cortisol responses and performance. This quadratic association might be predicted from the observation that moderately elevated cortisol responses may optimize cognitive functioning (Lupien et al., 2007). However, an inverted U-shaped association was not found, $b = 2.74$, $F(1, 222) = 0.01$, $p = .94$, 95% CI [-64.45, 69.94], $\eta_p^2 < .001$, and the quadratic term was dropped.

Table 4

Regression Coefficients Predicting Final Class Performance

Variable	Step 1				Step 2				Step 3			
	b	se(b)	CI(95%)		b	se(b)	CI(95%)		b	se(b)	CI(95%)	
Intercept	219.06***	3.51	212.15	225.97	39.18	25.22	-10.53	88.89	36.70	24.99	-12.55	85.95
Saliva Class 1 Cortisol ¹	-8.87	15.73	-39.86	22.12	-16.38	14.18	-44.32	11.56	-12.35	14.14	-40.12	15.51
Saliva Class 2-4 Cortisol ¹	19.88	17.66	-14.91	54.68	32.36 [†]	17.21	-1.56	66.27	42.40*	17.56	7.80	77.00
Race/ethnicity ²					-9.69**	3.62	-16.82	-2.55	-10.01**	3.59	-17.08	-2.93
Gender ³					-3.70	3.23	-10.06	2.66	-4.20	3.20	-10.51	2.11
SES ⁴					-0.22	4.09	-8.28	7.84	-0.05	4.05	-8.03	7.93
Prior performance					50.57***	6.69	37.38	63.76	50.89***	6.63	37.83	63.95
Missing prior performance ⁵					16.52	16.90	-16.79	49.83	22.42	16.92	-10.92	55.76
Time since awakening ⁶					0.53	1.88	-3.17	4.23	0.73	1.86	-2.93	4.40
Medication ⁷					7.77	10.73	-13.38	28.92	4.53	10.71	-16.58	25.64
Brush teeth ⁸					8.66	11.42	-13.84	31.15	8.94	11.30	-13.33	31.21
Cortisol X race/ethnicity									30.91*	13.09	5.11	56.70
<i>N</i> = 234	<i>R</i> ² = .006				<i>R</i> ² = .315***				<i>R</i> ² = .332**			

Note. *R*² change between steps significant at *p* < .001. ¹ Cortisol levels averaged then log-transformed and mean centered. ² Non-URM = -1, URM = 1. ³ Male = -1, Female = 1. ⁴

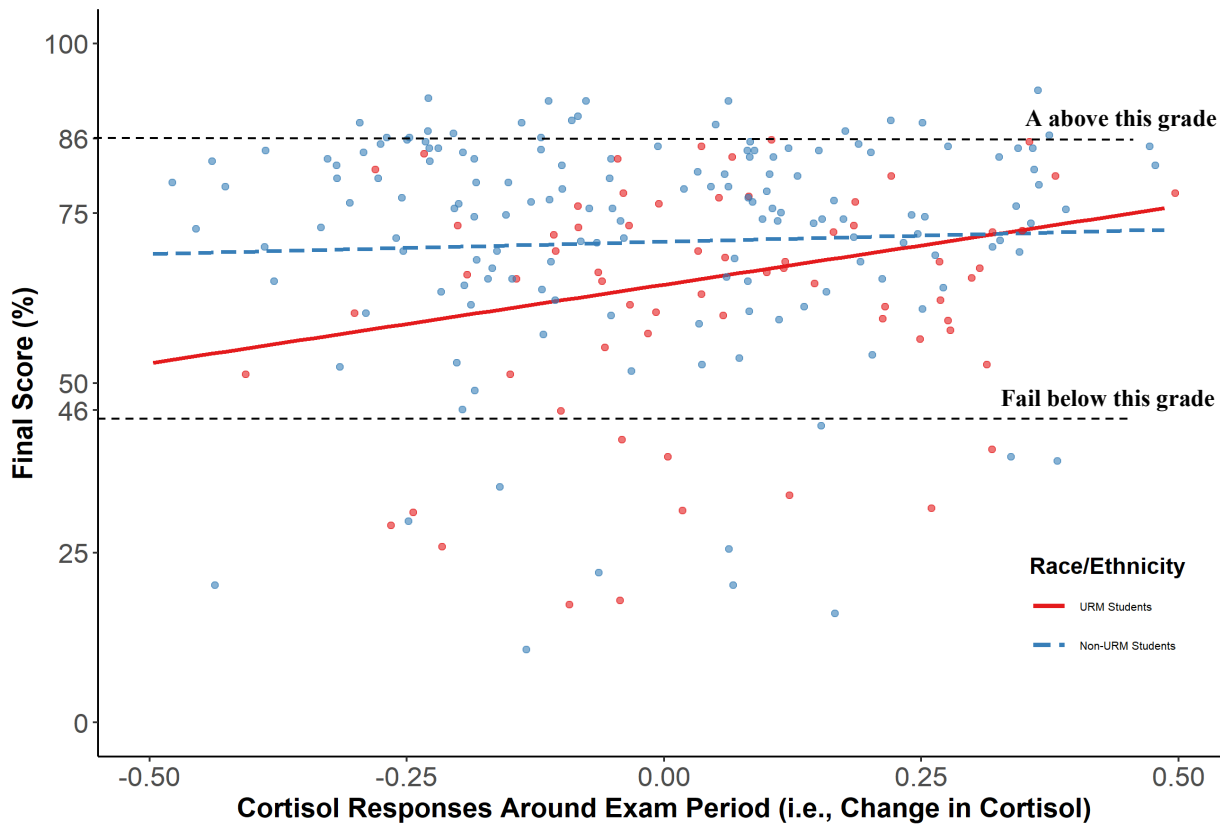
An average of standardized SES variables (parent’s income, parent’s education level, social status). ⁵ Participants with prior performance data = 0, Participants without prior performance data due to not consenting to release their transcript = 1. ⁶ Time since awakening averaged across four classes (in hours). ⁷ Whether students took medication 24 hours before saliva collection, averaged across the four classes in which saliva samples were collected. ⁸ Whether students had brushed their teeth an hour before the saliva collection, averaged across the four classes in which saliva samples were collected.

[†]*p* ≤ .10, **p* ≤ .05. ***p* ≤ .01. ****p* ≤ .001.

Step 3 revealed a significant interaction between cortisol responses and race/ethnicity in predicting final scores, $b = 30.91$, $F(1, 222) = 5.58$, $p = .02$, $\eta_p^2 = .025$ (see Figure 4 and Table 4). For URM students, cortisol responses predicted higher final scores, $b = 73.31$, $F(1, 222) = 9.09$, $p = .003$, 95% CI [25.40, 121.21], $\eta_p^2 = .039$. However, among non-URM students, cortisol responses were not associated with final scores, $b = 11.49$, $F(1, 222) = 0.36$, $p = .55$, 95% CI [-26.33, 49.31], $\eta_p^2 = .002$. Residual analyses revealed no outliers (see S20).

Figure 4

Students' Final Scores as a Function of Cortisol Responses Around the Exam Period and Race/Ethnicity



Note. The regression lines represent predicted final score in the class as a function of race/ethnicity for each observed value of cortisol around the exam period (i.e., averaged across Saliva classes 2-4), controlling for Saliva Class 1 cortisol and appropriate covariates (including Saliva Class 1 cortisol). Cortisol values have been log-transformed and mean centered.

Displayed points correspond to unadjusted observed data. Students who received over 86% received an A, and those who received under 46% failed the class; dotted lines illustrate the cutoffs for those grades.

Cortisol and Dropping the Course

To test the association of cortisol responses with whether students dropped the course (0 = completed, 1 = dropped), we used logistic regression. In *Step 1*, controlling for Saliva Class 1 cortisol, cortisol responses were not associated with whether students dropped the course, OR = .23, $\chi^2(1, N = 234) = 1.82, p = .18$ (see Table 5 for results from all steps and 95% CI). In *Step 2*, cortisol responses remained unassociated with dropping the course, OR = .17, $\chi^2(1, N = 234) = 1.81, p = .18$. However, *Step 3* revealed a marginal interaction between cortisol responses and race/ethnicity, OR = 0.14, $\chi^2(1, N = 234) = 3.42, p = .07$ (see Figure 5). Among URM students, higher cortisol responses significantly reduced the odds of dropping the course, OR = 0.03, $\chi^2(1, N = 234) = 4.11, p = .04, 95\% \text{ CI } [.001, 0.89]$. For non-URM students, cortisol responses were not associated with dropping the course, OR = 1.31, $\chi^2(1, N = 234) = 0.02, p = .88, 95\% \text{ CI } [.04, 41.91]$. Residual analyses revealed no outliers (see S20).

Table 5

Regression Coefficients Predicting Dropping the Course

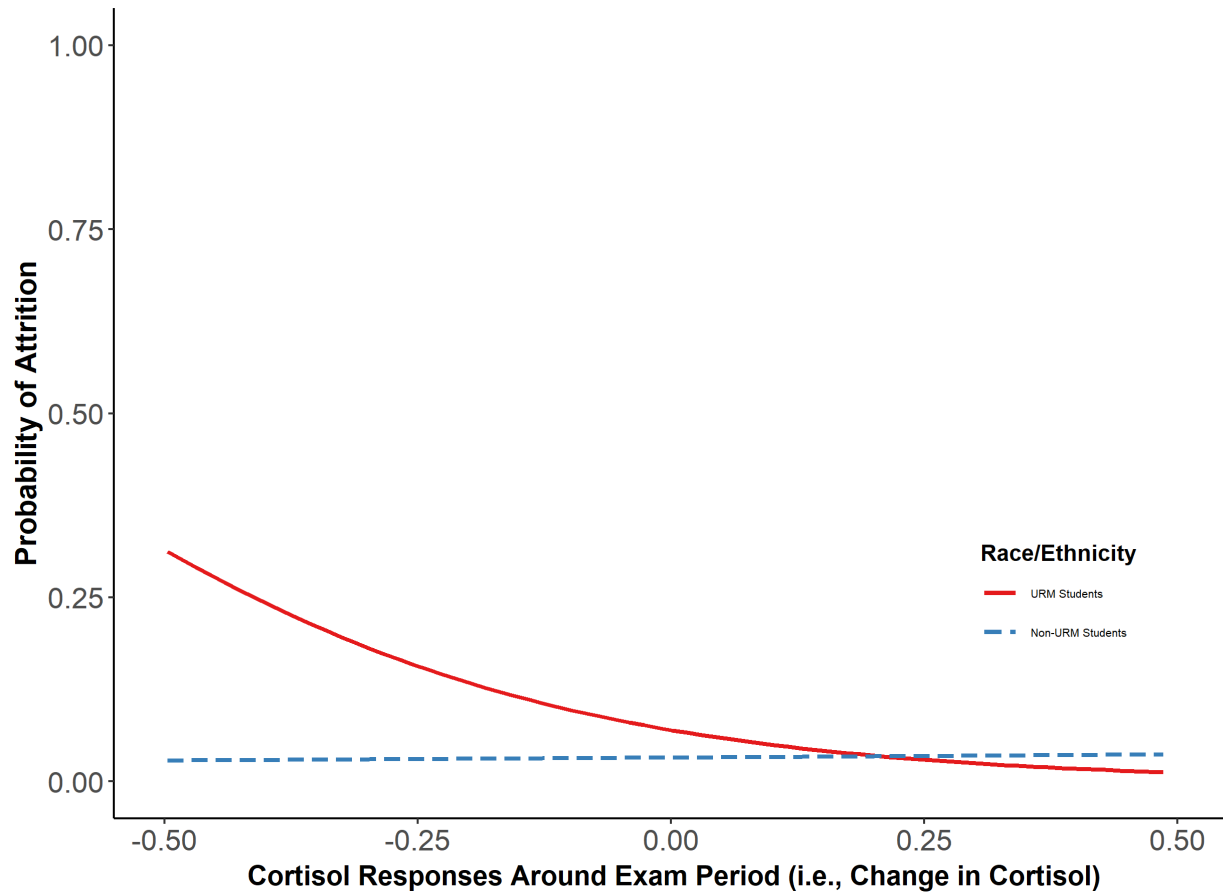
Variable	Step 1			Step 2			Step 3		
	Odds ratio	CI(95%)		Odds ratio	CI(95%)		Odds ratio	CI(95%)	
Intercept	0.11***			26.82†			34.09†		
Saliva Class 1 Cortisol ¹	2.29	0.33	15.99	2.53	0.30	21.54	1.51	0.16	14.28
Slave Class 2-4 Cortisol ¹	0.23	0.03	1.93	0.17	0.01	2.26	0.19	0.01	3.10
Race/ethnicity ²				1.60†	0.97	2.65	1.50	0.90	2.50
Gender ³				1.42	0.81	2.49	1.47	0.83	2.61
SES ⁴				1.90†	1.00	3.61	1.97*	1.04	3.74
Prior performance				0.24**	0.09	0.65	0.22**	0.08	0.62
Missing prior performance ⁵				0.00	0.00	0.00	0.00	0.00	0.00
Time since awakening ⁶				0.89	0.66	1.19	0.88	0.65	1.19
Medication ⁷				0.44	0.08	2.53	0.51	0.09	3.01
Brush teeth ⁸				0.16	0.01	1.79	0.15	0.02	1.55
Cortisol X race/ethnicity							0.14†	0.02	1.13
<i>N</i> = 234	Nagelkerke(<i>pseudo</i>) <i>R</i> ² = .02			Nagelkerke(<i>pseudo</i>) <i>R</i> ² = .196 *			Nagelkerke(<i>pseudo</i>) <i>R</i> ² = .226**		

¹ Cortisol levels averaged then log-transformed and mean centered. ² Non-URM = -1, URM = 1. ³ Male = -1, Female = 1. ⁴ An average of standardized SES variables (parent’s income, parent’s education level, social status). ⁵ Participants with prior performance data = 0, Participants without prior performance data due to not consenting to release their transcript = 1. ⁶ Time since awakening averaged across four classes (in hours). ⁷ Whether students took medication 24 hours before saliva collection, averaged across the four classes in which saliva samples were collected. ⁸ Whether students had brushed their teeth an hour before the saliva collection, averaged across the four classes in which saliva samples were collected.

†*p* ≤ .10, **p* ≤ .05. ***p* ≤ .01. ****p* ≤ .001.

Figure 5

Students' Probability of Dropping the Class as a Function of Cortisol Responses Around the Exam Period and Race/Ethnicity



Note. The regression lines represent the predicted probability of dropping the class as a function of race/ethnicity for each observed value of cortisol around the exam period (i.e., averaged across Saliva Classes 2-4), controlling for appropriate covariates (including Saliva Class 1 cortisol). Cortisol values have been log-transformed and mean centered.

Cortisol Responses and Students' Exam 1 Scores

In *Step 1*, controlling for Saliva Class 1 cortisol, students' Saliva Class 2 cortisol levels were not associated with their Exam 1 performance, $b = -2.02$, $F(1, 189) = 0.24$, $p = .63$, 95% CI $[-10.22, 6.18]$, $\eta_p^2 = .001$. In *Step 2*, cortisol responses were still not associated with Exam 1

performance, $b = 0.54$, $F(1, 181) = 0.02$, $p = .89$, 95% CI [-7.44, 8.51], $\eta_p^2 < .001$. In *Step 3*, a significant interaction emerged between students' cortisol responses and their race/ethnicity, $b = 6.61$, $F(1, 180) = 4.94$, $p = .03$, 95% CI [0.74, 12.47], $\eta_p^2 = .03$. For URM students, higher cortisol responses predicted marginally higher Exam 1 scores, $b = 10.44$, $F(1, 180) = 3.04$, $p = .08$, 95% CI [-1.37, 22.24], $\eta_p^2 = .017$, whereas among non-URM students, cortisol responses were not associated with Exam 1 scores, $b = -2.78$, $F(1, 180) = 0.42$, $p = .52$, 95% CI [-11.20, 5.64], $\eta_p^2 = .002$.

Exploratory Analyses - Why are Cortisol Responses Positively Associated with Performance for URM Students?

In this section, we explore why cortisol was positively associated with performance for URM students.

Cortisol and Self-Reported Focus and Engagement

Using data from the embedded smartphone study, we tested whether cortisol was positively associated with self-reported focus and engagement for URM students, which may help explain why cortisol positively predicted their academic performance. The 98 smartphone study participants were prompted to answer two relevant questions immediately only after Saliva Class 2 – 4 (but not Saliva Class 1) which overlapped with smartphone study timeline on a scale from 1 (*often*) to 4 (*never*): “How often were you focused in class today?” and “How often were you engaged in class today?” Both were reversed scored (i.e., higher scores indicate greater focus and engagement). These measures were completed 135 times by a subset of 76 students (42 URM) who also provided saliva samples during at least one of these classes. We used these measures as outcomes in two-level models with cortisol as the primary predictor, as described

below. Given the relatively small sample size, these analyses are likely under-powered, particularly to detect interaction effects.

Level-1 predictors were students' average (log-transformed) within-class cortisol levels for Saliva Class 2, 3, and 4, as well as the usual cortisol-related covariates. Level-2 predictors were students' URM status, as well as their gender and SES, as predictors of the Level-1 intercept, which was allowed to freely vary. All other Level-1 coefficients were fixed. We also included URM status as a predictor of the Level-1 cortisol slope. As elsewhere, we conducted theoretically motivated simple effects of the association between cortisol and focus/engagement as a function of URM status.¹ Below, we specify the multilevel models for these analyses. The subscript t = the class that cortisol was collected, and i = individuals.

$$\text{Level 1: Focus/Engagement}_{ti} = \beta_{0i} + \beta_{1i}(\text{Cortisol}_{ti}) + \beta_{2i}(\text{Time since awakening}_{ti}) + \beta_{3i}(\text{Medication}_{ti}) + \beta_{5i}(\text{Brush teeth}_{ti}) + e_{ti}$$

$$\text{Level 2: } \beta_{0i} = \gamma_{00i} + \gamma_{01i}(\text{Race}_i) + \gamma_{02i}(\text{SES}_i) + \gamma_{03i}(\text{Gender}_i) + \gamma_{04i}(\text{Race X Cortisol}_i) + u_{0i}$$

Higher cortisol levels predicted greater focus for URM students, $\beta = 1.12$, $t(54) = 2.10$, $p = .04$, 95% CI [0.09, 2.16], but not non-URM students, $\beta = 0.52$, $t(54) = 1.23$, $p = .23$, 95% CI [-0.30, 1.34], though the interaction between race/ethnicity and cortisol was not significant, $\beta = 0.30$, $t(54) = 1.06$, $p = .29$, 95% CI [-0.25, 0.86] (see Table S2). Cortisol did not predict self-reported engagement for either URM students, $\beta = 0.71$, $t(54) = 1.28$, $p = .21$, 95% CI [-0.37, 1.79], or non-URM students, $\beta = 0.48$, $t(54) = 1.06$, $p = .30$, 95% CI [-0.40, 1.37], but the

¹ We also exploratorily tested whether the interactions between cortisol responses and gender and SES significantly predicted measures of engagement. They did not (see S18).

direction of the association was positive and greater in magnitude for URM students. These analyses support the idea that higher cortisol was associated with more self-reported focus for URM students, but not the self-reported engagement question (though the direction was positive).

Association Between Social Identity Threat, Cortisol Responses, and Performance for URM Students

One potential reason that cortisol was associated with focus and better performance for URM students may be because URM students feel a need to strive to overcome negative stereotypes and succeed.

Ancillary analyses support this possibility. For URM students, those with higher baseline social identity threat perceived the class as more important, $b = 0.14$, $F(1, 267) = 6.25$, $p = .01$, 95% CI [0.03, 0.25], $\eta_p^2 = .023$. Social identity threat was not associated with perceived importance of the class for non-URM students (social identity threat x URM status interaction, $p = .01$). Assessing the class as more important may suggest a high motivation to succeed.

In line with the possibility noted above, in ancillary mediation models, results partially suggest that higher cortisol responses for URM students (but not for non-URM students) mediated the association between perceiving higher social identity threat and better performance outcomes (see S21 and Figure S2).

Discussion

Using data from a longitudinal field study in a gateway STEM course, we aimed to better understand cortisol response patterns in an academic setting and their association with academic performance. We were interested in determining whether cortisol responses and cortisol's association with performance differed by race/ethnicity.

Results suggest that the classroom setting and approaching exam did not acutely elicit cortisol responses. Contrary to our expectations, cortisol levels *declined* overall from the beginning to the end of each class and across classes from the beginning of the semester until after Exam 1. In retrospect, there are reasons to think that a traditional lecture-based class may not elevate cortisol even if the class was stressful. Cortisol reliably rises in situations that are socially evaluative, uncontrollable, and novel (Dickerson & Kemeny, 2004). Traditional lecture-based classes may not seem socially evaluative to students, who are mostly passive recipients of information. The absence of tests during typical classes may have further reduced any potential for the classroom setting to feel socially evaluative or uncontrollable. Also, the classroom could have lacked novelty because students became accustomed to the setting. Within classes, declining cortisol concentration may also, in part, be a function of time, since cortisol levels decline throughout the day (Kirschbaum & Hellhammer, 1989).

Students' cortisol levels may have not increased during the exam period because the anticipation of a written examination may similarly have lacked sufficient social evaluation to elevate cortisol since grading is temporally separated from the source of stress (i.e., the exam). As students become accustomed to their courses, classes around the exam period may not feel particularly more uncontrollable, novel, or socially-evaluative compared to other classes. In fact, an early-semester class may elicit a greater cortisol response because the course was still novel. Interestingly, cortisol trajectories did not track increasing self-reported stress, which increased as Exam 1 approached (assessed by smartphone study, see Figure S1). Future studies utilizing salivary cortisol in educational settings should consider the type of stress that may elicit cortisol responses.

Contrary to our pre-registered hypothesis, neither cortisol levels nor longitudinal patterns

of cortisol differed by race/ethnicity even though URM students performed worse in the class. Our results are among the first to investigate whether students' cortisol levels differed by race/ethnicity in an academic setting and contradict our expectation that URM students would have higher cortisol due to identity threat (Levy et al., 2016). This finding may suggest that cortisol did not capture stress in this setting.

Surprisingly, cortisol responses predicted better performance and greater persistence among URM students. In response to social identity threat URM students may either disengage to protect self-worth or strive to overcome negative stereotypes and adversity (James et al., 1983; Major & O'Brien, 2005). Among URM students, greater cortisol responses may represent movement from a relatively disengaged state to one characterized by active engagement and striving, in which higher cortisol is associated with greater learning and performance. Consistent with this idea, among URM students, greater reported social identity threat was associated with perceiving the class as more important, and higher cortisol predicted greater self-reported focus. This interpretation is also in line with research finding that high-performing URM students often strive to prove themselves in academic settings (Fries-Britt & Griffin, 2007). While elevated cortisol in the short-term may be beneficial for URM students' performance, over time it can have detrimental health consequences.

For non-URM students, cortisol responses were not associated with performance, leaving the implication of cortisol ambiguous. Non-URM students may have less reason than URM students to strive or disengage, and if cortisol did not capture stress, there may be little reason to expect it to predict performance. More research is needed to identify how the implication of cortisol can be altered by social psychological factors related to group membership.

Limitations & Future Directions

While it is a strength to have assessed cortisol repeatedly in class and to have used an early-semester baseline, we did not assess cortisol outside of class or on the day of the exam. The former could have served as an out-of-class baseline. However, finding a day that would represent a reliable low-stress baseline during students' busy academic calendars could be difficult. Without exam-day cortisol, we cannot know if cortisol levels would have been elevated on the exam day compared to other days, which might have interfered with cognition and memory retrieval. However, assessing cortisol on the day of the exam was not possible to avoid distracting students. If the exam was causing stress captured by cortisol, we would think that cortisol levels would have been elevated during Saliva Class 2, which immediately preceded the exam, relative to Saliva Class 1, but they were not.

Because of cortisol's diurnal pattern (Kirschbaum & Hellhammer, 1989), students' cortisol levels would naturally be higher in the morning than afternoon sessions of the class, suggesting that class time may be affecting results. To address this issue, we controlled for students' time since awakening and tested whether class time systemically changed the longitudinal pattern of cortisol and cortisol's effect on academic performance. It did not (see S6), helping alleviate the concern that class time was driving results.

While we controlled for a variety of relevant covariates, it is always possible that others we did not assess could have affected results. For example, sleep quality (Bassett et al., 2015) and diet and nutrition (Tomiya et al., 2010) can affect cortisol, but these were not available for us to control. Thus, results should be interpreted with respect to the important but non-exhaustive list of covariates included.

Future researchers interested in whether and how cortisol affects performance may benefit from assessing cortisol from hair instead of saliva. Hair samples can provide cumulative

cortisol level across weeks or months, providing a more stable assessment of cortisol than saliva samples.

Conclusion

The current study contributes to the understanding of the cortisol responses and their association with performance in a stressful real-life STEM academic setting. The present findings suggest that implications of cortisol can be affected by how people perceive and react to their environment. Cortisol in an academic setting may sometimes be associated with engagement and positively with performance, particularly for students who might be at risk of perceiving higher threat.

References

- Aardal, E., & Holm, A. C. (1995). Cortisol in saliva--reference ranges and relation to cortisol in serum. *European journal of clinical chemistry and clinical biochemistry : journal of the Forum of European Clinical Chemistry Societies*, 33(12), 927–932.
<https://doi.org/10.1515/cclm.1995.33.12.927>
- Adam, E., & Kumari, M. (2009). Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*, 34(10), 1423-1436.
<https://doi.org/10.1016/j.psyneuen.2009.06.011>.
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology*, 19(6), 586-592.
<https://doi.org/http://dx.doi.org/10.1037/0278-6133.19.6.586>
- Akinola, M., Fridman, I., Mor, S., Morris, M. W., & Crum, A. J. (2016). Adaptive Appraisals of Anxiety Moderate the Association between Cortisol Reactivity and Performance in Salary Negotiations. *PLoS One*, 11(12), e0167977.
<https://doi.org/10.1371/journal.pone.0167977>
- APA (2020). *Stress in America 2020. A national mental health crisis*. Washington, DC: American Psychological Association.
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Bassett, S. M., Lupis, S. B., Gianferante, D., Rohleder, N., & Wolf, J. M. (2015). Sleep quality but not sleep quantity effects on cortisol responses to acute psychosocial stress. *Stress*, 18(6), 638-644. <https://doi.org/10.3109/10253890.2015.1087503>

- Blascovich, J., & Mendes, W. B. (2010). Social Psychophysiology and Embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., pp. 194-227). NY: Wiley. <https://doi.org/10.1002/9780470561119.socpsy001006>
- Carnevale, A. P., Cheah, B., & Hanson, A. R. (2015). The economic values of college majors. Retrieved from <https://cew.georgetown.edu/wpcontent/uploads/Exec-Summary-web-B.pdf>
- Chen, E., Sheldon, C., & Miller, G. E. (2009). How Low Socioeconomic Status Affects 2-Year Hormonal Trajectories in Children. *Psychological Science*, *21*(1), 31-37. <https://doi.org/10.1177/0956797609355566>
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307-1310. <https://doi.org/10.1126/science.1128317>
- Cohen, G. L., & Garcia, J. (2005). "I am us": negative stereotypes as collective threats. *Journal of Personality and Social Psychology*, *89*(4), 566. <https://doi.org/10.1037/0022-3514.89.4.566>
- Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology*, *102*(3), 479. <https://doi.org/10.1037/a0026312>
- Crum, A. J., Akinola, M., Martin, A., & Fath, S. (2017). The role of stress mindset in shaping cognitive, emotional, and physiological responses to challenging and threatening stress. *Anxiety, Stress, & Coping*, *30*(4), 379-395. <https://doi.org/10.1080/10615806.2016.1275585>

- de Quervain, D. J., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience*, *3*, 313-314. <https://doi.org/10.1038/73873>
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355-391. <https://doi.org/10.1037/0033-2909.130.3.355>
- Dozier, M., Manni, M., Gordon, M. K., Peloso, E., Gunnar, M. R., Stovall-McClough, K. C., Eldreth, D., & Levine, S. (2006). Foster children's diurnal production of cortisol: An exploratory study. *Child maltreatment*, *11*(2), 189-197. <https://doi.org/10.1177/1077559505285779>
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, *13*(3), 478-492. <https://doi.org/10.1187/cbe.13-10-0204>
- Elzinga, B. M., & Roelofs, K. (2005). Cortisol-induced impairments of working memory require acute sympathetic activation. *Behavioral Neuroscience*, *119*(1), 98-103. <https://doi.org/10.1037/0735-7044.119.1.98>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. <https://doi.org/10.3758/BF03193146>
- Flinn, M. V. (2006). Evolution and ontogeny of stress response to social challenges in the human child. *Developmental Review*, *26*(2), 138-174. <https://doi.org/https://doi.org/10.1016/j.dr.2006.02.003>

- Fries-Britt, S., & Griffin, K. (2007). The Black box: How high-achieving Blacks resist stereotypes about Black Americans. *Journal of College Student Development*, 48(5), 509-524. <https://doi.org/10.1353/csd.2007.0048>
- Granger, D. A., Hibel, L. C., Fortunato, C. K., & Kapelewski, C. H. (2009). Medication effects on salivary cortisol: tactics and strategy to minimize impact in behavioral and developmental science. *Psychoneuroendocrinology*, 34(10), 1437-1448. <https://doi.org/10.1016/j.psyneuen.2009.06.017>
- Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways?. *PNAS Nexus*, 1(4), pgac167. <https://doi.org/10.1093/pnasnexus/pgac167>
- Heissel, J. A., Adam, E. K., Doleac, J. L., Figlio, D. N., & Meer, J. (2018). Testing, stress, and performance: How students respond physiologically to high-stakes testing. *Education Finance and Policy*, 1-50. https://doi.org/10.1162/edfp_a_00306
- Hostinar, C. E., McQuillan, M. T., Mirous, H. J., Grant, K. E., & Adam, E. K. (2014). Cortisol Responses to a Group Public Speaking Task for Adolescents: Variations by Age, Gender, and Race. *Psychoneuroendocrinology*, 50, 155-166. <https://doi.org/10.1016/j.psyneuen.2014.08.015>
- Hoyt, L. T., Zeiders, K. H., Ehrlich, K. B., & Adam, E. K. (2016). Positive upshots of cortisol in everyday life. *Emotion*, 16(4), 431-435. <https://doi.org/10.1037/emo0000174>
- James, S. A., Hartnett, S. A., & Kalsbeek, W. D. (1983). John Henryism and blood pressure differences among black men. *Journal of Behavioral Medicine*, 6(3), 259-278. <https://doi.org/10.1007/bf01315113>

- Jamieson, J. P., Black, A. E., Pelaia, L. E., & Reis, H. T. (2022). The impact of mathematics anxiety on stress appraisals, neuroendocrine responses, and academic performance in a community college sample. *Journal of Educational Psychology*.
<https://doi.org/10.1037/edu0000636>
- Keller, J., Bless, H., Blomann, F., & Kleinböhl, D. (2011). Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology*, 47(4), 849-852.
<https://doi.org/https://doi.org/10.1016/j.jesp.2011.02.004>
- King, R. B., McInerney, D. M., Ganotice, F. A., Jr, & Villarosa, J. B. (2015). Positive affect catalyzes academic engagement: Cross-sectional, longitudinal, and experimental evidence. *Learning and Individual Differences*, 39, 64-72.
<https://doi.org/https://doi.org/10.1016/j.lindif.2015.03.005>
- Kirschbaum, C., & Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology*, 22(3), 150-169. <https://doi.org/10.1159/000118611>
- Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262-265.
<https://doi.org/10.1126/science.1261375>
- Levy, D. J., Heissel, J. A., Richeson, J. A., & Adam, E. K. (2016). Psychological and biological responses to race-based social stress as pathways to disparities in educational outcomes. *American Psychologist*, 71(6), 455-473. <https://doi.org/10.1037/a0040322>
- Loft, P., Thomas, M. G., Petrie, K. J., Booth, R. J., Miles, J., & Vedhara, K. (2007). Examination stress results in altered cardiovascular responses to acute challenge and lower cortisol.

Psychoneuroendocrinology, 32(4), 367-375.

<https://doi.org/https://doi.org/10.1016/j.psyneuen.2007.02.004>

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition.

Brain and Cognition, 65(3), 209-237.

<https://doi.org/https://doi.org/10.1016/j.bandc.2007.02.007>

Major, B., & O'Brien, L. T. (2005). The Social Psychology of Stigma. *Annual Review of*

Psychology, 56(1), 393-421. <https://doi.org/10.1146/annurev.psych.56.091103.070137>

Mattarella-Micke, A., Mateo, J., Kozak, M. N., Foster, K., & Beilock, S. L. (2011). Choke or thrive? The relation between salivary cortisol and math performance depends on individual differences in working memory and math-anxiety. *Emotion*, 11(4), 1000.

<https://doi.org/10.1037/a0023224>

McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York academy of sciences*, 840(1), 33-44. <https://doi.org/10.1111/j.1749-6632.1998.tb09546.x>

McEwen, B. S. (2019). What Is the Confusion With Cortisol? *Chronic Stress*, 3,

247054701983364. <https://doi.org/10.1177/2470547019833647>

Michel, J. O., Campbell, C. M., & Dilsizian, K. (2018). Is STEM too Hard? Using Biglan to

Understand Academic Rigor and Teaching Practices across Disciplines. *Journal of the Professoriate*, 9(2).

Morgan, C. A., 3rd, Wang, S., Rasmusson, A., Hazlett, G., Anderson, G., & Charney, D. S.

(2001). Relationship among plasma cortisol, catecholamines, neuropeptide Y, and human

- performance during exposure to uncontrollable stress. *Psychosomatic Medicine*, 63(3), 412-422. <https://doi.org/10.1097/00006842-200105000-00010>
- National Center for Science and Engineering Statistics. (2021). *Women, minorities, and persons with disabilities in science and engineering: 2021*. Special Report NSF 21-321. National Science Foundation. <https://nces.nsf.gov/wmpd>
- Ng, V., Koh, D., & Chia, S. E. (2003). Examination stress, salivary cortisol, and academic performance. *Psychological Reports*, 93(3 Pt 2), 1133-1134. <https://doi.org/10.2466/pr0.2003.93.3f.1133>
- Payne, J. D., Jackson, E. D., Hoscheidt, S., Ryan, L., Jacobs, W. J., & Nadel, L. (2007). Stress administered prior to encoding impairs neutral but enhances emotional long-term episodic memories. *Learning & Memory*, 14(12), 861-868. <https://doi.org/10.1101/lm.743507>
- Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2017). R package "nlme" [Computer software]. Retrieved from <http://cran.r-project.org/package=nlme>.
- Pletzer, B., Wood, G., Moeller, K., Nuerk, H.-C., & Kerschbaum, H. H. (2010). Predictors of performance in a real-life statistics examination depend on the individual cortisol profile. *Biological Psychology*, 85(3), 410-416. <https://doi.org/https://doi.org/10.1016/j.biopsycho.2010.08.015>
- Preuß, D., Schoofs, D., Schlotz, W., & Wolf, O. T. (2010). The stressed student: Influence of written examinations and oral presentations on salivary cortisol concentrations in university students. *Stress*, 13(3), 221-229. <https://doi.org/10.3109/10253890903277579>
- R Core Team. (2020). R: A language and environment for statistical computing [Computer software]. Retrieved from <http://www.R-project.org/>.

- Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *International journal of STEM education*, 5(1). <https://doi.org/10.1186/s40594-018-0115-6>
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances* (pp. 91-116). Russel Sage Foundation.
- Rutkowski, K., Sowa, P., Rutkowska-Talipska, J., Kuryliszyn-Moskal, A., & Rutkowski, R. (2014). Dehydroepiandrosterone (DHEA): hypes and hopes. *Drugs*, 74(11), 1195-1207. <https://doi.org/https://doi.org/10.1007/s40265-014-0259-8>
- Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews*, 21(1), 55-89. <https://doi.org/10.1210/edrv.21.1.0389>
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379-399. <https://doi.org/https://doi.org/10.1037/h0046234>
- Schoofs, D., Hartmann, R., & Wolf, O. T. (2008). Neuroendocrine stress responses to an oral academic examination: No strong influence of sex, repeated participation and personality traits. *Stress: The International Journal on the Biology of Stress*, 11(1), 52-61. <https://doi.org/10.1080/10253890701453943>

- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613-629. <https://doi.org/10.1037/0003-066X.52.6.613>
- Tomiyama, A. J., Mann, T., Vinas, D., Hunger, J. M., DeJager, J., & Taylor, S. E. (2010). Low calorie dieting increases cortisol. *Psychosomatic medicine*, 72(4), 357. [10.1097/PSY.0b013e3181d9523c](https://doi.org/10.1097/PSY.0b013e3181d9523c)
- Turetsky, K. M., Purdie-Greenaway, V., Cook, J. E., Curley, J. P., & Cohen, G. L. (2020). A psychological intervention strengthens students' peer social networks and promotes persistence in STEM. *Science Advances*, 6(45), eaba9221. <https://doi.org/10.1126/sciadv.aba9221>
- Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: implications for the classroom. *NPJ Science Of Learning*, 1, 16011. <https://doi.org/10.1038/npjscilearn.2016.11>
- Whitcomb, K. M., & Singh, C. (2021). Underrepresented minority students receive lower grades and have higher rates of attrition across STEM disciplines: A sign of inequity?. *International Journal of Science Education*, 43(7), 1054-1089. <https://doi.org/10.1080/09500693.2021.1900623>
- Yerkes, R.M., Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*. 18 (5): 459–482. [10.1002/cne.920180503](https://doi.org/10.1002/cne.920180503)